

PTI

V. Poirier  
 RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/471,255

DATE: 03/21/2001  
 TIME: 15:27:37

Input Set : A:\ES.txt  
 Output Set: N:\CRF3\03212001\I471255.raw

ENTERED

4 <110> APPLICANT: BIOCHEM PHARMA INC.  
 5 HAMEL, Jos,e  
 6 BRODEUR, Bernard R.  
 7 PINEAU, Isabelle  
 8 MARTIN, Denis  
 9 RIOUX, Cl,ment  
 11 <120> TITLE OF INVENTION: NOVEL STREPTOCOCCUS ANTIGENS  
 14 <130> FILE REFERENCE: 12806-11PCT  
 C--> 16 <140> CURRENT APPLICATION NUMBER: US/09/471,255  
 C--> 16 <141> CURRENT FILING DATE: 1999-12-23  
 16 <150> PRIOR APPLICATION NUMBER: US 60/113,800  
 17 <151> PRIOR FILING DATE: 1998-12-23  
 19 <160> NUMBER OF SEQ ID NOS: 102  
 21 <170> SOFTWARE: FastSEQ for Windows Version 3.0  
 23 <210> SEQ ID NO: 1  
 24 <211> LENGTH: 3120  
 25 <212> TYPE: DNA  
 26 <213> ORGANISM: S. pneumoniae  
 28 <220> FEATURE:  
 31 <400> SEQUENCE: 1  
 32 atg aaa ttt agt aaa aaa tat ata gca gct gga tca gct gtt atc gta 48  
 33 tcc ttg agt cta tgt gcc tat gca cta aac cag cat cgt tcg cag gaa 96  
 34 aat aag gac aat aat cgt gtc tct tat gtg gat ggc agc cag tca agt 144  
 35 cag aaa agt gaa aac ttg aca cca gac cag gtt agc cag aaa gaa gga 192  
 36 att cag gct gag caa att gta atc aaa att aca gat cag ggc tat gta 240  
 37 acg tca cac ggt gac cac tat cat tac tat aat ggg aaa gtt cct tat 288  
 38 gat gcc ctc ttt agt gaa gaa ctc ttg atg aag gat cca aac tat caa 336  
 39 ctt aaa gac gct gat att gtc aat gaa gtc aag ggt ggt tat atc atc 384  
 40 aag gtc gat gga aaa tat tat gtc tac ctg aaa gat gca gct cat gct 432  
 41 gat aat gtt cga act aaa gat gaa atc aat cgt caa aaa caa gaa cat 480  
 42 gtc aaa gat aat gag aag gtt aac tct aat gtt gtc gta gca agg tct 528  
 43 cag gga cga tat acg aca aat gat ggt tat gtc ttt aat cca gct gat 576  
 44 att atc gaa gat acg ggt aat gct tat atc gtc cct cat gga ggt cac 624  
 45 tat cac tac att ccc aaa agc gat tta tct gtc gtc gaa tta gca gca 672  
 46 gct aaa gca cat ctg gtc ggt gga aaa aat atg caa ccg agt cag tta agc 720  
 47 tat tct tca aca gct aat gac aat aac acg caa tct gta gca aaa gga 768  
 48 tca act acg aag cca gca aat aaa tct gaa aat ctc cag agt ctt ttg 816  
 49 aag gaa ctc tat gat tca cct acg gcc caa cgt tac agt gaa tca gat 864  
 50 ggc ctg gtc ttt gac cct gtc aat ggt att atc acg aca cca aat gga 912  
 51 gtt gcg att ccg cat ggc gac cat tac cac ttt att cct tac acg aag 960  
 52 ctt tct gtc tta gaa gaa aat ggc aga atg gtc cct atc agt gga 1008  
 53 act ggt tct aca gtt tct aca aat gca aaa cct aat gaa gta gtc tct 1056  
 54 agt cta ggc agt ctt tca acg aat cct tct ttc tta acg aca agt aag 1104  
 55 gag ctc tct tca gca tct gat ggt tat att ttt aat cca aaa gat atc 1152  
 56 gtt gaa gaa acg gct aca gct tat att gta aga cat ggt gat cat ttc 1190  
 57 cat tac att cca aaa tca aat caa att ggg caa ccg act ctt cca aac 1248  
 58 aat agt cta gca aca cct tct cca tct ctt cca atc aat cca gga act 1296

RECEIVED  
 MAR 28 2001  
 TECH CENTER 1600/2900

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/471,255

DATE: 03/21/2001

TIME: 15:27:37

Input Set : A:\ES.txt

Output Set: N:\CRF3\03212001\I471255.raw

59	tca cat gag aaa cat gaa gaa gat gga tac gga ttt gat gct aat cgt	1344
60	att atc gct gaa gat gaa tca ggt ttt gtc atg agt cac gga gac cac	1392
61	aat cat tat ttc ttc aag aag gac ttg aca gaa gag caa att aag gct	1440
62	gcg caa aaa cat tta gag gaa gtt aaa act agt cat aat gga tta gat	1488
63	tct ttg tca tct cat gaa cag gat tat cca ggt aat gcc aaa gaa atg	1536
64	aaa gat tta gat aaa aaa atc gaa gaa aaa att gct ggc att atg aaa	1584
65	caa tat ggt gtc aaa cgt gaa agt att gtc gtg aat aaa gaa aaa aat	1632
66	gcg att att tat ccg cat gga gat cac cat cat gca gat ccg att gat	1680
67	gaa cat aaa ccg gtt gga att ggt cat tct cac agt aac tat gaa ctg	1728
68	ttt aaa ccc gaa gaa gga gtt gct aaa aaa gaa ggg aat aaa gtt tat	1776
69	act gga gaa gaa tta acg aat gtt gtt aat ttg tta aaa aat agt acg	1824
70	ttt aat aat caa aac ttt act cta gcc aat ggt caa aaa cgc gtt tct	1872
71	ttt agt ttt ccg cct gaa ttg gag aaa aaa tta ggt atc aat atg cta	1920
72	gta aaa tta ata aca cca gat gga aaa gta ttg gag aaa gta tct ggt	1968
73	aaa gta ttt gga gaa gga gta ggg aat att gca aac ttt gaa tta gat	2016
74	caa cct tat tta cca gga caa aca ttt aag tat act atc gct tca aaa	2064
75	gat tat cca gaa gta agt tat gat ggt aca ttt aca gtt cca acc tct	2112
76	tta gct tac aaa atg gcc agt caa acg att ttc tat cct ttc cat gca	2160
77	ggg gat act tat tta aga gtg aac cct caa ttt gca gtg cct aaa gga	2208
78	act gat gct tta gtc aga gtg ttt gat gaa ttt cat gga aat gct tat	2256
79	tta gaa aat aac tat aaa gtt ggt gaa atc aaa tta ccg att ccg aaa	2304
80	tta aac caa gga aca acc aga acg gcc gga aat aaa att cct gta acc	2352
81	ttc atg gca aat gct tat ttg gac aat caa tcg act tat att gtg gaa	2400
82	gta cct atc ttg gaa aaa gaa aat caa act gat aaa cca agt att cta	2448
83	cca caa ttt aaa agg aat aaa gca caa gaa aac tca aaa ctt gat gaa	2496
84	aag gta gaa gaa cca aag act agt gag aag gta gaa aaa gaa aaa ctt	2544
85	tct gaa act ggg aat agt act agt aat tca acg tta gaa gaa gtt cct	2592
86	aca gtg gat cct gta caa gaa aaa gta gca aaa ttt gct gaa agt tat	2640
87	ggg atg aag cta gaa aat gtc ttg ttt aat atg gac gga aca att gaa	2688
88	tta tat tta cca tca gga gaa gtc att aaa aag aat atg gca gat ttt	2736
89	aca gga gaa gca cct caa gga aat ggt gaa aat aaa cca tct gaa aat	2784
90	gga aaa gta tct act gga aca gtt gag aac caa cca aca gaa aat aaa	2832
91	cca gca gat tct tta cca gag gca cca aac gaa aaa cct gta aaa cca	2880
92	gaa aac tca acg gat aat gga atg ttg aat cca gaa ggg aat gtg ggg	2928
93	agt gac cct atg tta gat cca gca tta gag gaa gct cca gca gta gat	2976
94	cct gta caa gaa aaa tta gaa aaa ttt aca gct agt tac gga tta ggc	3024
95	tta gat agt gtt ata ttc aat atg gat gga acg att gaa tta aga ttg	3072
96	cca agt gga gaa gtg ata aaa aag aat tta tct gat ttc ata gcg	3117
97	taa	3120
99	<210> SEQ ID NO: 2	
100	<211> LENGTH: 1039	
101	<212> TYPE: PRT	
102	<213> ORGANISM: S. pneumoniae	
104	<400> SEQUENCE: 2	
105	Met Lys Phe Ser Lys Lys Tyr Ile Ala Ala Gly Ser Ala Val Ile Val	
106	1 5 10 15	
107	Ser Leu Ser Leu Cys Ala Tyr Ala Leu Asn Gln His Arg Ser Gln Glu	
108	20 25 30	
109	Asn Lys Asp Asn Asn Arg Val Ser Tyr Val Asp Gly Ser Gln Ser Ser	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/471,255

DATE: 03/21/2001

TIME: 15:27:37

Input Set : A:\ES.txt

Output Set: N:\CRF3\03212001\I471255.raw

110	35	40	45	
111	Gln Lys Ser Glu Asn Leu Thr Pro Asp Gln Val Ser Gln Lys Glu Gly			
112	50	55	60	
113	Ile Gln Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val			
114	65	70	75	80
115	Thr Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr			
116	85	90	95	
117	Asp Ala Leu Phe Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Gln			
118	100	105	110	
119	Leu Lys Asp Ala Asp Ile Val Asn Glu Val Lys Gly Gly Tyr Ile Ile			
120	115	120	125	
121	Lys Val Asp Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala			
122	130	135	140	
123	Asp Asn Val Arg Thr Lys Asp Glu Ile Asn Arg Gln Lys Gln Glu His			
124	145	150	155	160
125	Val Lys Asp Asn Glu Lys Val Asn Ser Asn Val Ala Val Ala Arg Ser			
126	165	170	175	
127	Gln Gly Arg Tyr Thr Thr Asn Asp Gly Tyr Val Phe Asn Pro Ala Asp			
128	180	185	190	
129	Ile Ile Glu Asp Thr Gly Asn Ala Tyr Ile Val Pro His Gly Gly His			
130	195	200	205	
131	Tyr His Tyr Ile Pro Lys Ser Asp Leu Ser Ala Ser Glu Leu Ala Ala			
132	210	215	220	
133	Ala Lys Ala His Leu Ala Gly Lys Asn Met Gln Pro Ser Gln Leu Ser			
134	225	230	235	240
135	Tyr Ser Ser Thr Ala Ser Asp Asn Asn Thr Gln Ser Val Ala Lys Gly			
136	245	250	255	
137	Ser Thr Ser Lys Pro Ala Asn Lys Ser Glu Asn Leu Gln Ser Leu Leu			
138	260	265	270	
139	Lys Glu Leu Tyr Asp Ser Pro Ser Ala Gln Arg Tyr Ser Glu Ser Asp			
140	275	280	285	
141	Gly Leu Val Phe Asp Pro Ala Lys Ile Ile Ser Arg Thr Pro Asn Gly			
142	290	295	300	
143	Val Ala Ile Pro His Gly Asp His Tyr His Phe Ile Pro Tyr Ser Lys			
144	305	310	315	320
145	Leu Ser Ala Leu Glu Glu Lys Ile Ala Arg Met Val Pro Ile Ser Gly			
146	325	330	335	
147	Thr Gly Ser Thr Val Ser Thr Asn Ala Lys Pro Asn Glu Val Val Ser			
148	340	345	350	
149	Ser Leu Gly Ser Leu Ser Ser Asn Pro Ser Ser Leu Thr Thr Ser Lys			
150	355	360	365	
151	Glu Leu Ser Ser Ala Ser Asp Gly Tyr Ile Phe Asn Pro Lys Asp Ile			
152	370	375	380	
153	Val Glu Glu Thr Ala Thr Ala Tyr Ile Val Arg His Gly Asp His Phe			
154	385	390	395	400
155	His Tyr Ile Pro Lys Ser Asn Gln Ile Gly Gln Pro Thr Leu Pro Asn			
156	405	410	415	
157	Asn Ser Leu Ala Thr Pro Ser Pro Ser Leu Pro Ile Asn Pro Gly Thr			
158	420	425	430	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/471,255

DATE: 03/21/2001

TIME: 15:27:37

Input Set : A:\ES.txt

Output Set: N:\CRF3\03212001\I471255.raw

159 Ser His Glu Lys His Glu Glu Asp Gly Tyr Gly Phe Asp Ala Asn Arg  
 160 435 440 445  
 161 Ile Ile Ala Glu Asp Glu Ser Gly Phe Val Met Ser His Gly Asp His  
 162 450 455 460  
 163 Asn His Tyr Phe Phe Lys Lys Asp Leu Thr Glu Glu Gln Ile Lys Ala  
 164 465 470 475 480  
 165 Ala Gln Lys His Leu Glu Glu Val Lys Thr Ser His Asn Gly Leu Asp  
 166 485 490 495  
 167 Ser Leu Ser Ser His Glu Gln Asp Tyr Pro Gly Asn Ala Lys Glu Met  
 168 500 505 510  
 169 Lys Asp Leu Asp Lys Lys Ile Glu Glu Lys Ile Ala Gly Ile Met Lys  
 170 515 520 525  
 171 Gln Tyr Gly Val Lys Arg Glu Ser Ile Val Val Asn Lys Glu Lys Asn  
 172 530 535 540  
 173 Ala Ile Ile Tyr Pro His Gly Asp His His Ala Asp Pro Ile Asp  
 174 545 550 555 560  
 175 Glu His Lys Pro Val Gly Ile Gly His Ser His Ser Asn Tyr Glu Leu  
 176 565 570 575  
 177 Phe Lys Pro Glu Glu Gly Val Ala Lys Lys Glu Gly Asn Lys Val Tyr  
 178 580 585 590  
 179 Thr Gly Glu Glu Leu Thr Asn Val Val Asn Leu Leu Lys Asn Ser Thr  
 180 595 600 605  
 181 Phe Asn Asn Gln Asn Phe Thr Leu Ala Asn Gly Gln Lys Arg Val Ser  
 182 610 615 620  
 183 Phe Ser Phe Pro Pro Glu Leu Glu Lys Lys Leu Gly Ile Asn Met Leu  
 184 625 630 635 640  
 185 Val Lys Leu Ile Thr Pro Asp Gly Lys Val Leu Glu Lys Val Ser Gly  
 186 645 650 655  
 187 Lys Val Phe Gly Glu Gly Val Gly Asn Ile Ala Asn Phe Glu Leu Asp  
 188 660 665 670  
 189 Gln Pro Tyr Leu Pro Gly Gln Thr Phe Lys Tyr Thr Ile Ala Ser Lys  
 190 675 680 685  
 191 Asp Tyr Pro Glu Val Ser Tyr Asp Gly Thr Phe Thr Val Pro Thr Ser  
 192 690 695 700  
 193 Leu Ala Tyr Lys Met Ala Ser Gln Thr Ile Phe Tyr Pro Phe His Ala  
 194 705 710 715 720  
 195 Gly Asp Thr Tyr Leu Arg Val Asn Pro Gln Phe Ala Val Pro Lys Gly  
 196 725 730 735  
 197 Thr Asp Ala Leu Val Arg Val Phe Asp Glu Phe His Gly Asn Ala Tyr  
 198 740 745 750  
 199 Leu Glu Asn Asn Tyr Lys Val Gly Glu Ile Lys Leu Pro Ile Pro Lys  
 200 755 760 765  
 201 Leu Asn Gln Gly Thr Thr Arg Thr Ala Gly Asn Lys Ile Pro Val Thr  
 202 770 775 780  
 203 Phe Met Ala Asn Ala Tyr Leu Asp Asn Gln Ser Thr Tyr Ile Val Glu  
 204 785 790 795 800  
 205 Val Pro Ile Leu Glu Lys Glu Asn Gln Thr Asp Lys Pro Ser Ile Leu  
 206 805 810 815  
 207 Pro Gln Phe Lys Arg Asn Lys Ala Gln Glu Asn Ser Lys Leu Asp Glu

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/471,255

DATE: 03/21/2001  
TIME: 15:27:37

Input Set : A:\ES.txt  
Output Set: N:\CRF3\03212001\I471255.raw

208	820	825	830
209	Lys Val Glu Glu Pro Lys Thr Ser Glu Lys Val Glu Lys Glu Lys Leu		
210	835	840	845
211	Ser Glu Thr Gly Asn Ser Thr Ser Asn Ser Thr Leu Glu Glu Val Pro		
212	850	855	860
213	Thr Val Asp Pro Val Gln Glu Lys Val Ala Lys Phe Ala Glu Ser Tyr		
214	865	870	875
215	Gly Met Lys Leu Glu Asn Val Leu Phe Asn Met Asp Gly Thr Ile Glu		
216	885	890	895
217	Leu Tyr Leu Pro Ser Gly Glu Val Ile Lys Lys Asn Met Ala Asp Phe		
218	900	905	910
219	Thr Gly Glu Ala Pro Gln Gly Asn Gly Glu Asn Lys Pro Ser Glu Asn		
220	915	920	925
221	Gly Lys Val Ser Thr Gly Thr Val Glu Asn Gln Pro Thr Glu Asn Lys		
222	930	935	940
223	Pro Ala Asp Ser Leu Pro Glu Ala Pro Asn Glu Lys Pro Val Lys Pro		
224	945	950	955
225	Glu Asn Ser Thr Asp Asn Gly Met Leu Asn Pro Glu Gly Asn Val Gly		
226	965	970	975
227	Ser Asp Pro Met Leu Asp Pro Ala Leu Glu Glu Ala Pro Ala Val Asp		
228	980	985	990
229	Pro Val Gln Glu Lys Leu Glu Lys Phe Thr Ala Ser Tyr Gly Leu Gly		
230	995	1000	1005
231	Leu Asp Ser Val Ile Phe Asn Met Asp Gly Thr Ile Glu Leu Arg Leu		
232	1010	1015	1020
233	Pro Ser Gly Glu Val Ile Lys Lys Asn Leu Ser Asp Phe Ile Ala		
234	1025	1030	1035
236	<210> SEQ ID NO: 3		
237	<211> LENGTH: 2523		
238	<212> TYPE: DNA		
239	<213> ORGANISM: S. pneumoniae		
241	<220> FEATURE:		
242	<221> NAME/KEY: CDS		
243	<222> LOCATION: (1)...(2520)		
244	<223> OTHER INFORMATION: Coding region of BVH-11 gene		
246	<400> SEQUENCE: 3		
247	atg aaa atc aat aaa aaa tat cta gct ggg tca gta gct aca ctt gtt		48
248	Met Lys Ile Asn Lys Lys Tyr Leu Ala Gly Ser Val Ala Thr Leu Val		
249	1 5 10 15		
251	tta agt gtc tgt gct tat gaa cta ggt ttg cat caa gct caa act gta		96
252	Leu Ser Val Cys Ala Tyr Glu Leu Gly Leu His Gln Ala Gln Thr Val		
253	20 25 30		
255	aaa gaa aat aat cgt gtt tcc tat ata gat gga aaa caa gcg acg caa		144
256	Lys Glu Asn Asn Arg Val Ser Tyr Ile Asp Gly Lys Gln Ala Thr Gln		
257	35 40 45		
259	aaa acg gag aat ttg act cct gat gag gtt acg aag cgt gaa gga atc		192
260	Lys Thr Glu Asn Leu Thr Pro Asp Glu Val Ser Lys Arg Glu Gly Ile		
261	50 55 60		
263	aac gcc gaa caa atc gtc atc aag att acg gat caa ggt tat gtg acc		240

**VERIFICATION SUMMARY** DATE: 03/21/2001  
**PATENT APPLICATION:** US/09/471,255 TIME: 15:27:38

Input Set : A:\ES.txt  
Output Set: N:\CRF3\03212001\I471255.raw

L:16 M:270 C: Current Application Number differs, Replaced Current Application No  
L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date